

U.S. DEPARTMENT OF LABOR



# Enterprise Data Strategy



# FOREWORD

Let me thank the amazing public servants who helped craft this inaugural data strategy for the Department of Labor, in particular our Chief Data Officer Scott Gibbons and the entire team at our Office of Data Governance, along with our Chief Innovation Officer Chike Aguh.

Since I came to the department, I have asked my team to unleash their full power—to take every tool, our statutory authority, rules and regulations, funding, and the incredible talent we have across the country and unleash it every single day for the good of America’s working people.

Too often, an assumption in government is, *I need a change in the law to do this, or the only way to make a real impact is with more money, or if I could hire more people, we could make this happen*, instead of focusing on the tremendous power we already have and the many ways we are not fully utilizing it in service to our mission. This includes examining all the ways we have tied our own hands, interpreted policies and authorities too narrowly, been too afraid to try a new way of operating. In assessing organizational risk, I ask that we examine not just the risk of certain activities—which is important—but the risk of not acting.

Data is truly one of the superpowers of the Department of Labor. That superpower allows us to move the needle on critical goals for American workers:

- **Job Quality:** Since the start of the Biden-Harris administration, it has been clear that one of the most important tasks before us is getting Americans back to work (or into work for the first time) after the pandemic. We resolved to help provide them with not just any jobs, but quality jobs that pay family sustaining wages; ensure security, dignity and respect; and provide a path to the American Dream. This is why the department worked to create the first-of-its-kind Job Quality Measurement Initiative with various foundations to define and advance job quality at scale across our economy.
- **Equity:** As evidenced by President Biden signing his Executive Order on Equity on day one and by our department’s equity strategy, we are committed to ensuring that every American worker is protected, respected and advancing in this economy. To do this, we are lifting up the needs and voices of workers from communities who have been overlooked: workers of color, immigrants, women, workers with disabilities, those living in concentrated poverty, and those in rural and/or low-income communities. We do this not simply because it is the right thing to do but because building an inclusive economy measured by the experiences of these workers is the way to make sure no one is left behind. Data is critical to fulfilling our aims here. In addition to planning our work through the collection and evaluation of reliable data, we are committed to making data available. The Bureau of Labor Statistics is now publishing monthly jobs data on Native American and Native Alaskan workers and working on disaggregating demographic data for the nation’s Asian American, Native Hawaiian and Pacific Islander communities, which are diverse and have different needs and economic outcomes.

This data strategy builds on these examples and many others happening all over the department. Per our strategy, we will make our data findable by anyone, accessible to all, interoperable across siloes, and reusable across time. Beyond that, we will work to find ways to put this data into the hands of workers themselves, to empower them to make decisions that will improve their lives.

Ending where I began, I want to thank all those who worked to bring this strategy to life, and I want to offer even deeper thanks to those who will use this data to improve life for all of us. We will unleash the full power of data at the Department of Labor and leverage it on behalf of the workers we serve. Thank you.

---

**Julie Su**  
*Deputy Secretary of Labor*



## EXECUTIVE SUMMARY

This data strategy guides current and prospective efforts across the Department of Labor (DOL) to create and improve capabilities that support better alignment of data operations, more effective data management, and greater use of data to inform program administration and decision-making. Aligning disparate data definitions, collections, and management processes to allow their products to integrate and function effectively is a challenge, but one with clear solutions that require a unified approach to improving data, people, culture, analytics, and technology. By identifying clear and realistic goals that are expected to bring value by harmonizing and optimizing data operations, working to improve data governance through objective measures of progress and value, and improving data production, DOL builds capabilities for data to inform and support decision making.

This three-year strategy builds off the foundation of missions and laws that guide the Department and is composed of four guiding principles and five goals that can develop more consistent and effective data governance and align Departmental and agency planning to improve data management and use capabilities. The four guiding principles are (1) making data findable, (2) making data accessible, (3) making data interoperable, and (4) making data reusable; or the FAIR principles for short. The five strategic goals are: (1) ensuring data are managed to be open by default, (2) ensuring data are comprehensible, (3) ensuring data are fit for purpose, (4) ensuring data are available in consistent and predictable ways, and (5) managing data as an enterprise asset, incorporate stakeholder input, and as appropriate, made public in ways that provide real benefit to data users. Taken together, the foundation, principles, and goals form a strategy that should guide decision-making, project planning, and data governance, and incrementally improve the state of data at DOL. The strategy concludes with focused areas for improvement at the agency and Departmental levels including (1) ensuring data quality, (2) building and maintaining data talent, (3) integration of data into existing agency management and planning systems to create a practical and realizable path forward, and (4) expanding the data capabilities for producing sophisticated analytics.

# BACKGROUND FOR THE DOL DATA STRATEGY

DOL is composed of agencies with varying missions, programs, and organizational structures, with each producing data in response to diverse mandates, laws, and policies. Creating data that can serve as a strategic asset across such an enterprise requires a strong focus on the quality, availability, and utility of such data. Data products that are consistent in format and content increase value by making it easier to find patterns and trends that are relevant to program administration, and reduce the costs and complexity of analysis through efficient and predictable use. High-quality data are essential to understanding the effectiveness of DOL programs, the beneficiaries and stakeholders impacted by those programs, and the degree to which we meet our missions.

Numerous factors shape the manner and extent to which Federal agencies leverage data as a strategic asset: the laws and policies governing data they generate or use; the missions, methods, and stakeholders; the capabilities in the data infrastructure; and their organizational data culture. Data culture is a particularly important factor, encompassing how the agency perceives, defines, collects, and manages data, and the data acumen and skills among staff to leverage data.

## Legal and Policy Context

Recent laws and policies covering Federal data management<sup>1</sup> have challenged Federal agencies to evaluate and improve how data are treated as a strategic asset and made broadly available to stakeholders. Federal agencies are reconsidering the role that data can play in guiding the planning, operations, budget, and management of programs. Federal policies, including executive orders expanding the role for data in supporting

<sup>1</sup> Including but not limited to the Foundations for Evidence-Based Policymaking Act of 2018, the Geospatial Data Act of 2018, the Grant Reporting Efficiency and Agreements Transparency Act of 2019, and the Digital Accountability and Transparency Act of 2014

artificial intelligence<sup>2</sup>, scientific integrity<sup>3</sup>, and equity<sup>4</sup>, along with the Federal Data Strategy<sup>5</sup>, similarly seek to align agencies in the ways they seek to build data capacity. The complexity of this agenda necessitates that Federal agencies assess their current data infrastructure and capacity, identify areas to improve and align, and develop methods to mature their ability to define, collect, manage, use, and share data.

## Assessments of the Current State of Data

During fiscal year (FY) 2020, DOL's Data Board conducted two assessments of DOL's infrastructure, capacities, and processes. The first was a data management maturity model assessment of capabilities including data management, human capital, data culture, analytics, and technological capabilities. The assessment ranked DOL's capabilities as developing, with the lowest scores in the broad areas of human capital and technology. The second was a data skills assessment, fielded within a maturity model framework that looked at a variety of capabilities fundamental to effective data management and use. The skills assessment identified specific areas to focus on for improvement including building data talent to better access, visualize, and analyze data. During 2022, DOL should release its first evidence capacity assessment describing DOL's ability to generate and use scientific information including statistics, evaluation, research, and analysis to inform policy, planning, and program administration. These assessments have shown that while DOL is rich in data assets, it has additional work to do to bring the full potential of those assets to the stakeholders that we serve. The assessments provide an objective benchmark of strengths and weaknesses to guide planning and strategy, and ensure DOL can target actions to remediate known weaknesses and provide tangible and necessary improvements.

<sup>2</sup> <https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence>

<sup>3</sup> <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/memorandum-on-restoring-trust-in-government-through-scientific-integrity-and-evidence-based-policymaking/>

<sup>4</sup> <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/>

<sup>5</sup> <https://strategy.data.gov/>

---

***“As the Department’s Evidence Officials, we recognize the vital role this strategy plays in promoting evidence-based decision-making. Increasingly, our success in accomplishing the Department’s mission is dependent on recognizing the essential role data has in understanding the populations we serve, the current and emerging challenges those populations face, and providing an objective basis for efficiently and effectively targeting our resources and efforts to those most in need. Data are central to assessing the state of workers and the workforce, the impacts of Department of Labor programs, and providing transparency into our actions and the impacts that result. This Enterprise Data Strategy brings a more central focus to the need for quality, consistency, and availability of data to inform and influence how DOL carries out its mission. This strategy will guide efforts to improve Federal data management in predictable and consistent ways, and best position us to leverage those improvements to produce evidence that can guide strategic planning and program administration, inform policy and support data-informed decisions.”***

---

— Department of Labor Evidence Act Officials: Chief Evaluation Officer Dr. Christina Yancey, Chief Data Officer Mr. Scott Gibbons, BLS Commissioner of Labor Statistics Dr. William Beach, and Director of the Performance Management Center Mr. Dennis Johnson



## Stakeholders

DOL data are a product of interactions between our programs and the public. Strategic planning for data changes involves parties who play an important role across the lifecycle of data. Notable stakeholders who can provide input toward improving DOL data include:

- DOL programs staff and leadership, who define and collect statistical, administrative, budget, grant, and transactional data, and use it to guide program administration;
- External governmental, public and private entities who depend on DOL data to support their work, support collaborative efforts and the identification of patterns and trends that optimize data informed decision-making;
- American workers and members of the public who are covered by DOL regulations, jointly administer DOL programs, or are recipients of services originating from DOL programs;
- DOL's Chief Information Officer (CIO) and related Information Technology (IT) operations staff, who match needs with services, implement common

enterprise data management, and maintain evolving capabilities;

- DOL's senior leadership, who provide vision and priorities for Departmental programs, use data to monitor program impact and performance on stakeholders and beneficiaries, and support sound resource allocation; and
- DOL's data governance body and network of data stewards, who promote the value of and advocate for more optimal data governance, and ensure that data management and governance processes stay practically aligned with supporting agency mission and public benefit.

Data inform efforts to make operations appropriate, efficient, and effective, and to measure the impact the programs have. Stakeholder support for the ideas and actions in this strategy is a necessary pre-requisite to ensure that changes in data management across the enterprise are implemented successfully, with a strong customer focus, and in a manner that returns maximum value to the Department as a whole and the people who rely on its programs.



# FRAMEWORK FOR THE DATA STRATEGY

Data are essential inputs to virtually all management and administrative functions in Federal agencies. The quality of outputs from mission-critical tasks — such as strategic planning, analytics, assessment and performance management, evidence generation, enforcement, targeting activities, and resource allocation — is dependent on the quality of the data used to generate them. DOL’s data strategy identifies a series of principles and goals meant to increase the benefit from data operations; align agency and departmental data, IT, and analytical projects; and ensure that data can consistently and predictably meet the evolving needs of governmental, public, and private data users.

## Principles Guiding DOL Data Governance

To optimize all data-dependent operations at DOL, the Department must commit to abiding by the FAIR principles. These principles serve as pillars in the data strategy, and effectively function as a “pole star” that harmonizes and guides disparate planning, decision-making, and operations across the enterprise. Broad adoption of these principles will help to ensure that data are defined, collected, managed, and leveraged effectively, and made public in ways that best serve the needs of DOL and all its stakeholders.

**Principle #1: Data Must Be Findable.** For DOL to be a data-informed agency, data and associated metadata must be findable in ways that facilitate use by both humans and computers.

**Principle #2: Data Must Be Accessible.** For DOL to effectively leverage data to inform program administration, clear instructions for accessing all data assets must be provided, including authentication or authorization for any restricted-use data.

**Principle #3: Data Must Be Interoperable.** For DOL program data to be maximally useful across the Department, data and metadata must abide by common formats and standards.

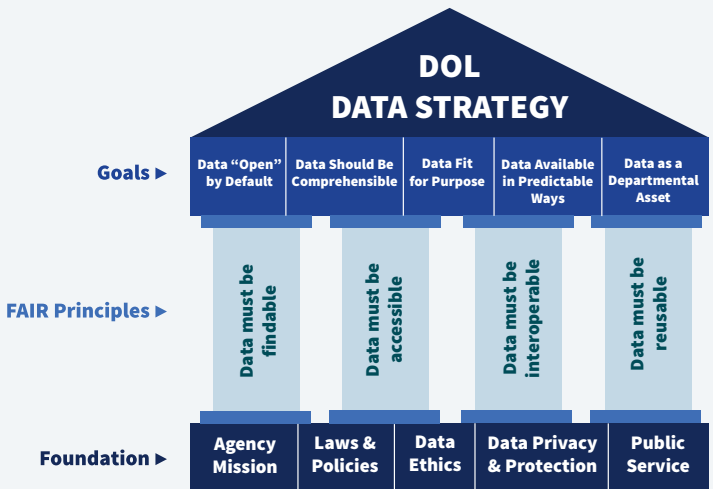


Fig. 1: Schematic of the Data Strategy Components

**Principle #4: Data Must Be Reusable.** For DOL data to support the broadest possible set of analytical use cases, data and metadata must be well documented and described so that they can be replicated or combined in different settings.

By committing to the FAIR principles, the DOL puts itself in a position to direct the necessary activities and acquire the needed capabilities to begin to achieve the strategic goals.

## Goals for Improved Data Governance

Growing data capacity and maturity within an enterprise requires a commitment to cultural, organizational, and technological change. A transition towards more effective data management and use requires DOL to address areas for improvement identified through assessment; develop systems, processes, and people who can support a more effective data infrastructure; and maintain a consistent focus on the central importance of data and the value it holds.

# FAIR PRINCIPLES



Fig. 2: The FAIR Principles



**Goal #1: Data Should be Considered “Open” by Default**

DOL data should be developed and managed to be open by default. DOL should continue to maintain and expand comprehensive inventories of data assets to ensure people have awareness of the data DOL produces and manages. Inventories should serve as a clearinghouse of supporting information about the data including content, metadata, and names of data stewards and custodians to support data users. Inventories and the data described in them should be openly accessible to the greatest extent possible, in accordance with privacy and security laws. When data cannot be made publicly available, DOL programs should consider offering the data through restricted use access programs or produce an explanation for restricting access to the data. To foster openness and transparency, DOL should publish an enterprise data inventory of all its data assets that includes any restrictions on data access.

**Goal #2: Data Should Be Comprehensible**

Comprehensible data are those for which descriptions of the data design, meaning, structure, format, and provenance are available. Comprehensible data can be correctly understood by those users who played no role in the creation of the data. This understanding is provided through appropriate documentation of the data. Without this documentation, a data set is a collection of letters and numerals without any context for interpreting meaning within the data. With appropriate documentation, it is possible to understand what the data content signify, how they are produced, how they can be used, and what restrictions there are to their use. Documentation provides a functional and operational framework to interpret and use the data like a Rosetta Stone: without it, data cannot be properly understood; with it, data are made comprehensible.

**Goal #3: Data Should Be Fit for Purpose**

Data are fit for purpose when any ambiguity about their content, format, origin, and appropriate use is eliminated. This requires evaluating several factors, including but not limited to documentation, format, coverage, design, quality, and access. DOL should strive to deliver data that meet this criterion.

**Goal #4: Data Should Be Readily Available in Consistent and Predictable Ways**

DOL should establish “data-as-a-service” infrastructure, providing relevant subsets of data, on-demand, and in a machine-readable format, allowing for direct integration of these data into business intelligence, analytical software and visualization systems. DOL should also create shared services for data access and data transfer, so data can flow from sources to services in predictable and efficient ways. DOL should transition siloed data systems towards common storage frameworks, ensuring DOL staff can leverage these services and capabilities with appropriate governance, and efficiently establish data quality frameworks. Finally, DOL should develop a consolidated data portal that simplifies data access, provides a customer-focused experience, and offers resources that allow people to confidently and independently locate and access the data they need.

**Goal #5: Data Should Be a Departmental Strategic Asset**

For data to effectively function as a departmental strategic asset, DOL should define, manage, and collect data in ways that support the broadest possible set of use cases, and maximize the possibility for reuse within the enterprise. DOL should move to create and adopt common data structures and schemas that increase the value of data while decreasing the costs of data management. DOL should develop controlled templates for common data elements and agencies should transition towards using those data to create interoperability, reduce analysis burden, and support linkage of records. DOL data should conform to expected norms in content and format. DOL programs need to transition away from data silos and parochial data management and embrace the idea that data become a strategic asset when they are shared, exposed to emerging analytical use cases, and leveraged in a variety of ways to yield the greatest possible benefit.

# OPERATIONALIZING THE STRATEGY

DOL should undertake actions in a variety of areas to improve data governance, build data capacity, expand and improve data talent, and grow the ways in which data can inform and support the mission. Adopting the FAIR principles as key to planning and decision-making should help to ensure that even disparate efforts are aligned towards common goals. DOL programs should further their alignment with this strategy by committing to collaborating through the DOL Data Board and the Office of Data Governance (ODG) to work across organizational lines, and support enterprise approaches to data management, stewardship, architecture, and utilization. Initial areas of focus should be on organizational and cultural change, strengthening governance, increasing data talent, improving data documentation, modernizing data infrastructure, integrating data management into existing systems, and expanding data use to inform program administration.

## Organizational and Culture Change

Implementing this strategy begins with a culture change that places focus on the value data offer as a departmental strategic asset. DOL agencies should commit to collaborating with the DOL Data Board, DOL's CIO, and the ODG to define and implement common definitions, formats, and quality criteria for common data elements to ensure interoperability of data. As DOL agencies collect, define, manage, and share data using more standardized and optimal approaches and for the broadest possible set of analytical use cases, DOL data should be fit for purpose and have greater value to inform decision-making and evidence building.

## Strengthened Governance

The Department should strengthen data governance by creating formal, defined standards for common data elements to make data predictable and consistent, and implementing them wherever practicable. In addition, DOL should continue to develop the current network of data stewards by offering training and socializing the concepts of more optimal data management, and demonstrating the value proposition and cost efficiencies that come from harmonizing data definitions, management approaches, and use patterns. A particularly important prerequisite for improving governance, analytics, and open data is creating consistent and effective documentation associated with managing DOL data sets. DOL should work through the DOL Data Board to promote wider use of appropriate data standards, create effective metadata, and improve data documentation. Documentation should be available in plain language, offer common descriptions of the meaning of data, and expand descriptions of data format and organization.

## Developing, Recruiting, and Maintaining Data Talent

A critical part of DOL's Data Strategy is to create a workforce within DOL that is educated and trained on the theory, techniques, and infrastructure for effective data management and innovative analytical uses of data. To achieve this goal, DOL should define and formalize standardized roles and position requirements for those data-centric occupations, align the provisioning of tools with the skill sets of data scientists and data professionals, develop clarity around DOL's analytical strategy, improve planning for the development of data teams within agencies, and create formal career development paths for data personnel. DOL should work to retain effective data talent. Attrition among data staff produces tremendous inefficiencies and leads to time spent recruiting, hiring, and training, instead of benefitting from completed data science work. Offering productive environments that inspire data talent to produce innovative analytical products, and promoting complementary efforts to stand up communities of practice, analytical seminars, opportunities to present analytical work to governmental, public, and private gatherings, and cross-program analysis opportunities can improve the prospects for retaining data talent and fostering a productive environment for data science to thrive.

---

**“As Chief Information Officer, the OCIO is in support of DOL’s Data Strategy. It aligns with OCIO’s goal to get the right data, to the right people, at the right time in a secure manner. OCIO provides enterprise-wide data and analytics services and is building a modern data platform for DOL’s agencies to advance mission and drive innovation. The DOL Data Strategy, coupled with leadership engagement and an enabling operational plan, will help OCIO accelerate the centralization and integration of DOL data to achieve operational efficiencies and improved insights in support of evidenced based decision making and alignment with DOL modernization activities.”**

---

**— Department of Labor Chief Information Officer Gundeep Ahluwalia**



## Improving the Quality of Departmental Data

Data—essential to setting policy and delivering efficiently and effectively on the Department’s mission of serving America’s workforce—are often not defined, collected, or managed in ways that best serve the needs of DOL or its stakeholders. There needs to be a more central focus on the quality, consistency, and availability of data to inform and influence how DOL carries out its mission. Data have good quality when they accurately represent events, characteristics, activities, or outcomes they record. DOL agencies should work through the Data Board and collaborate with the CIO, Chief Data Officer, and other governmental programs and agencies to define and implement data quality standards and practices to ensure DOL data are fit for purpose and interoperable. Uniform standards and practices for managing data are essential for establishing, maintaining, and improving data quality throughout the Department.

## Integrating Data as a Priority into Existing Management Systems

Improving data governance and data management is a complex process that often requires integration with organizational, budgeting, operational, personnel, and technological systems and processes. Institutionalizing changes requires integrating them into existing processes for strategic planning, resource allocation, and establishing priorities for staffing and IT. DOL will establish a process to have agencies integrate agency-specific and Departmental data projects they are undertaking to better manage and leverage data into their annual budget operating and Agency Management

Plans. Adding data projects to Agency Management Plans explicitly links improved data management and use with agency mission, and integrates data into existing frameworks for budget planning and management, IT priority setting, staffing priorities, and performance management systems with metrics and milestones for the work to be done. This approach avoids creation of new administrative process and generates efficient and effective integration of data into well-understood and closely monitored management systems. DOL will consolidate those agency plans into an annual data plan that includes both agency and Departmental priorities, and make this plan available to the public on an annual basis.

## Expanding Analytical Capabilities to Leverage Data to Inform Planning and Administration

DOL should develop a plan for enhancing and expanding the tools and capabilities that support exploration, analysis, visualization, and prediction systems using data. These tools are essential to DOL’s ability to effectively leverage data, recruit and retain data talent to support our program offices, and promote innovation in the ways data are used to develop evidence and program-relevant insights. In combination with improved data quality and methods for data access, improved analytical capabilities should help DOL create products that turn data into information and insights that more optimally guide planning, actions, and resource allocation. Increasing the quality and quantity of analytical work also reinforces the strategic value and importance of data to the people who collect and manage it, and serves as valuable feedback.

# PUTTING STRATEGY INTO PRACTICE

A number of specific and ongoing actions illustrate DOL's commitment to making progress towards the principles and goals identified in this strategy, and are designed to bring benefit to the administration of Federal programs as well as to generate improved data and data products for all of DOL's stakeholders, including:

**First-ever Open Data Request for Information (RFI):**

To understand the needs of workers and the myriad of stakeholders who consume our data, DOL will issue its first-ever Open Data RFI to solicit comments, feedback, and suggestions from the public for targeted improvements on data quality, availability, and modes of access.

**Further Develop Workforce Scorecards and Dashboards:** DOL continues to leverage data to create and explore more optimal formats and content

for workforce scorecards describing the efficacy of workforce development programs across the country. These tools will generate more optimal outcomes for numerous stakeholders in employment and training programs; from workers seeking the best-fit for training, to businesses looking for the best sources of talent, to providers working to make their services more effective and efficient.

**Combine and Improve Enforcement Data for Public Consumption:** Based on recommendations from the Worker Empowerment and Organizing Taskforce, DOL will create more interoperable and integrated data products by linking enforcement data across programs so the public can understand how companies are meeting their obligations under law.

**Working Across Stakeholders to Modernize the LMI System:** The Labor Market Information ecosystem across the country is a vital source of information for everyone from workers to policymakers, especially during shocks such as the recent pandemic. DOL seeks to bolster and strengthen this system and the partners inside and outside of government that support it, with the goal of making the entire system more responsive and more transparent around the state of all workers.



## SUMMARY

Modernizing perspectives on the value of data and increasing the sophistication of DOL data operations requires a comprehensive approach that includes working with data, people, culture, analytics, and technology. DOL should take a comprehensive approach to reframe the role of data that includes reinforcing the importance of designing for quality, accessibility, and utility in infrastructure, ensuring the quality and utility of data are appropriately central to emerging priorities, and maintaining focus on who we serve with our data.

Whether the intent is to optimize internal program operations, provide transparency into the impacts and beneficiaries of governmental programs, or create greater value for all public, private, and governmental data users, DOL should place a more central focus on ensuring that data provide maximum value, can be conveniently found, comprehensively understood, and confidently used. Findable, accessible, interoperable, and reusable data promote the delivery of efficient and effective services to the public, advances the management of Government information as an asset, and fuels entrepreneurship, innovation, and scientific discovery that optimally supports administration of DOL programs, meets agency mission, and brings benefit to DOL's core customers.

